Ann Inst Stat Math (2011) 63:1103–1122
DOI 10.1007/s10463-010-0284-7

Testing separability in marked multidimensional point processes with covariates

Chien-Hsun Chang · Frederic Paik Schoenberg

Received: 1 August 2008 / Revised: 30 March 2009 / Published online: 2 June 2010

Abstract In modeling marked point processes, it is convenient to assume a separable or multiplicative form for the conditional intensity, as this assumption typically allows one to estimate each component of the model individually. Tests have been proposed in the simple marked point process case, to investigate whether the mark distribution is separable from the spatial–temporal characteristics of the point process. Here, we extend these tests to the case of a marked point process with covariates, and where one is interested in testing the separability of each of the covariates, as well as the mark and the coordinates of the point process. The extension is not at all trivial, and covariates must be treated in a fundamentally different way than marks and coordinates of the process, especially when the covariates are not uniformly distributed. An application is given to point process models for forecasting wildfire hazard in Los Angeles County, California, and solutions are proposed to the problem of how to proceed when the separability hypothesis is rejected.

Keywords Conditional intensity · Covariates · Marked point processes · Separability · Testing · Wildfires

1 Introduction

In a spatial–temporal marked point process, it is typically assumed that marks are separable from the spatial–temporal process. The separability hypothesis is especially

C.-H. Chang · F. P. Schoenberg (✉)
Department of Statistics, University of California, 8125 Math-Science Building,
Los Angeles 90095-1554, USA
e-mail: frederic@stat.ucla.edu

C.-H. Chang
e-mail: cchang@stat.ucla.edu; tete_chang@hotmail.com

convenient since each component of a separable process may be modeled and estimated individually. For instance, in separable models commonly fit to the space–time locations and sizes of earthquakes, the earthquake size distribution is posited to be static over time and space, and thus may be modeled and estimated separately, without regard to the times and locations of earthquakes (Ogata 1988, 1998; Kagan and Jackson 1994). The separability assumption should, however, be tested before fitting a spatial–temporal marked point process (Assunção and Maia 2007).

Schoenberg (2004) and Assunção and Maia (2007) proposed several nonparametric test statistics to investigate the separability of spatial–temporal marked point processes for models without covariates. The purpose of this article is to extend these tests to the cases where the conditional intensity may depend not only on the location in question but also on time-varying covariates, where the parameters governing this dependence are to be estimated. An application is provided to models for forecasting wildfire hazard in Los Angeles County, using covariates such as temperature, relative humidity, precipitation, and wind speed.

Section 2 describes the Los Angeles County weather and wildfire data applied in this article. Section 3 reviews the definition of separability in spatial–temporal marked point processes and the separability tests outlined in Schoenberg (2004) and Assunção and Maia (2007). Methods for estimating the conditional intensity of models with covariates are presented and applied to wildfire hazard models in Sect. 4. Section 5 proposes solutions to the problem of how to proceed when the separability hypothesis is rejected, and this too is applied to wildfire hazard modeling in Los Angeles County. A summary and conclusions are given in Sect. 6.

2 Data

Since 1976, the National Fire Danger Rating System (NFDRS) has constructed various wildfire hazard indices, based on daily measurements of meteorological variables, which are in turn used by Fire Departments to prepare personnel and machinery and to plan wildfire suppression and prevention activities (see Pyne et al. 1996; Andrews and Bradshaw 1997). The inputs used to construct these ratings include a variety of meteorological variables, including air temperature, relative humidity, precipitation, wind speed, and wind direction, which are recorded at Remote Automatic Weather Stations (RAWS) across the United States (Bradshaw et al. 1983; Warren and Vance 1981). Some alternatives to the NFDRS indices, based on space–time marked point process models using the RAWS data as covariates, have been explored in Schoenberg et al. (2007, 2009).

We focus here on data from 16 RAWS located within Los Angeles County, California. Daily summaries of these measurements are collected each afternoon at 1 p.m. and transmitted via satellite to a central station for archiving. Detailed data on Los Angeles County wildfires have been collected and compiled by several agencies, including the Los Angeles County Fire Department (LACFD), the Los Angeles County Department of Public Works (LACDPW), the Santa Monica Mountains National Recreation Area, and the California Department of Forestry and Fire Protection. These wildfire data include the origin dates, the centroid locations, and the polygons mapping

Fig. 1 Centroid locations of 513 wildfires of at least 0.0405 km^2 (10 acres) observed in Los Angeles County from January 1976 to December 2000

the area burned in wildfires dating back to 1878. According to the LACFD, the wildfire data before 1950 is believed to be complete for fires burning at least 0.405 km^2 (100 acres), and though the LACFD has been mapping fires as small as 0.00405 km^2 (1 acre) since 1950, the data appear to be complete only for fires burning at least 0.0405 km^2 (10 acres) (Schoengerg et al. 2003).

In this article, we restrict our attention to temperature, relative humidity, precipitation, wind speed, and wildfires of at least 0.0405 km^2 (10 acres) recorded between January 1976 and December 2000. There were 592 wildfires of at least 0.0405 km^2 (10 acres) recorded in the time range considered; however, only 513 wildfires occurred on days with complete weather information. Figure 1 shows the centroid locations of these 513 wildfires. For further detail, including images of the spatial locations of these wildfires, and discussions of errors and missing values, see Peng et al. (2005) and Schoenberg et al. (2007).

3 Separability of spatial–temporal marked point processes

3.1 Spatial–temporal marked point processes

A point process N may be thought of as a random collection of points in some space $\mathcal{X}$. See Daley and Vere-Jones (2003) for a thorough treatment of point processes and their theoretical properties. For the case of spatial–temporal marked point processes, each point may be represented as a portion $(t, x, m) \in R^{n+2}$, where t is a 1-dimensional temporal coordinate, $x = (x_1, \ldots, x_n)$ is an n-dimensional spatial coordinate,

and m is a 1-dimensional mark. For instance, in modeling wildfire occurrences or earthquakes, each event is identified as a point with t the time of the event's origin, x the 2- or 3-dimensional location associated with the event (e.g. its centroid or its estimated location of origin), and m a real-valued measure of its size. For further examples see Schoenberg et al. (2002) and the references therein.

3.2 Separability

Marked point processes are typically modeled by specifying the conditional intensity function (CIF) of the process. When it exists, the conditional intensity $\lambda(t, x, m|H_t)$ may be defined as the limiting expected rate of occurrence of points per space–time-mark volume conditional on the history of the process prior to time t,

$$\lambda(t, x, m|H_t) = \lim_{|B|\to 0} \frac{E[N(B)|H_t]}{|B|}, \tag{1}$$

where $B = (t, t + \triangle t) \times (x, x + \triangle x) \times (m, m + \triangle m) \in R^{n+2}$, $N(B)$ is the total number of points in B, $|B|$ is the volume of B, and H_t is the history of the process prior to time t.

Following Cressie (1993) and Schoenberg (2004), we say the marked point process is *separable* with respect to the mark m if its CIF can be expressed as

$$\lambda(t, x, m|H_t) = \lambda_1(t, x|H_t) f(m), \tag{2}$$

where λ_1 is a nonnegative predictable process and f is a fixed nonnegative function. If the CIF may further be reduced to the form

$$\lambda(t, x, m|H_t) = \lambda_1(t|H_t) f_1(x) f_2(m), \tag{3}$$

where λ_1 is a nonnegative predictable process and f_1 and f_2 are fixed nonnegative functions, then we call the process *completely separable*. Completely separable models are typically too restrictive for real applications, but separability of the mark is commonly assumed in modeling spatial–temporal marked point processes.

3.3 Separability tests for marks

Before proceeding to the more general case of marked point processes with covariates, we briefly review previous work on tests for the simpler case of testing the separability of the mark, for a marked point process without covariates. Suppose that $\hat{\lambda}(t, x, m|H_t)$ is a nonparametric kernel estimate of the CIF of the spatial–temporal-marked point process N:

$$\hat{\lambda}(t, x, m|H_t) = \int_{\mathcal{X}} K_{n+2}(t - u, x - y, m - m')\mathrm{d}N(u, y, m'), \tag{4}$$

where here and in what follows we use the notation K_d to denote a d-dimensional kernel density. $\hat{\lambda}(t, x, m|H_t)$ may be considered a non-separable estimate of the CIF.

Let $\tilde{\lambda}_1(t, x|H_t)$ denote a nonparametric kernel estimate of the CIF of the marginal spatial–temporal point process consisting exclusively of the locations and times of points of N and ignoring the marks:

$$\tilde{\lambda}_1(t, x|H_t) = \int_{\mathcal{X}} K_{n+1}(t-u, x-y)\mathrm{d}N(u, y, m). \tag{5}$$

Similarly, consider a nonparametric kernel estimate $\tilde{f}(m)$ of the mark density f,

$$\tilde{f}(m|H_t) = \frac{1}{N(\mathcal{X})} \int_{\mathcal{X}} K_1(m-m')\mathrm{d}N(t, x, m'), \tag{6}$$

where $N(\mathcal{X})$ is the total number of points in $\mathcal{X}$. Under the null hypothesis of separability (2), a separable estimate of $\lambda(t, x, m|H_t)$ is given by

$$\tilde{\lambda}(t, x, m|H_t) = \tilde{\lambda}_1(t, x|H_t)\tilde{f}(m). \tag{7}$$

In what follows, the terms H_t will be dropped in order to simplify the notation.

For details and guidelines for optimally selecting densities and bandwidths for kernel smoothing, and for correcting for boundary effects, see Stone (1984), Silverman (1986), Sheather and Jones (1991), Bowman and Azzalini (1997), and Venables and Ripley (2002).

Under the null hypothesis of separability (2), the two CIF estimates $\hat{\lambda}(t, x, m)$ and $\tilde{\lambda}(t, x, m)$ should be similar. Schoenberg (2004) proposed several nonparametric test statistics:

$$S_1 = \sup\left\{\left|\hat{\lambda}(t, x, m) - \tilde{\lambda}(t, x, m)\right| \Big/ \sqrt{\tilde{\lambda}(t, x, m)} \; ; (t, x, m) \in \mathcal{X}\right\},$$

$$S_2 = \inf\left\{\left|\hat{\lambda}(t, x, m) - \tilde{\lambda}(t, x, m)\right| \Big/ \sqrt{\tilde{\lambda}(t, x, m)} \; ; (t, x, m) \in \mathcal{X}\right\},$$

$$S_3 = \int_0^T \int_{R^n} \int_R \left[\hat{\lambda}(t, x, m) - \tilde{\lambda}(t, x, m)\right]^2 \mathrm{d}m \, \mathrm{d}x \, \mathrm{d}t,$$

$$S_4 = \int_{\mathcal{X}} \left\{\log\left[\hat{\lambda}(t, x, m)\right] - \log\left[\tilde{\lambda}(t, x, m)\right]\right\} \mathrm{d}N$$
$$- \int_0^T \int_{R^n} \int_R \left[\hat{\lambda}(t, x, m) - \tilde{\lambda}(t, x, m)\right] \mathrm{d}m \, \mathrm{d}x \, \mathrm{d}t,$$

$$S_5 = \frac{1}{N(\mathcal{X})} \sum_i \left[\hat{\lambda}(t_i, x_i, m_i) - \tilde{\lambda}(t_i, x_i, m_i)\right]^2,$$

$$S_6 = \max\left[\hat{\lambda}(t, x, m) - \tilde{\lambda}(t, x, m)\right]^2.$$

Large values of any of these test statistics suggest departures from separability. Schoenberg (2004) used simulated data generated by different types of models to

investigate the performance of these test statistics, and found that S_3 was very powerful at detecting gradual changes in the distribution of the marks though not as powerful in detecting clustering or inhibition in the marks. As suggested in Schoenberg (2004), p-values for these test statistics may readily be obtained using simulations of separable marked point processes each with CIF equal to $\tilde{\lambda}(t, x, m)$.

Assunção and Maia (2007) derived a score test statistic T and showed that there is a close relationship between the score test statistic T and the test statistics S_4 and S_5 proposed by Schoenberg (2004). They assumed that, if the process is non-separable, there exists a constant ϵ and a certain predictable function $g(t, x, m)$ such that

$$\lambda(t, x, m) = \lambda_1(t, x) f(m)[1 + \epsilon g(t, x, m)], \tag{8}$$

where $\epsilon g(t, x, m)$ is the relative difference between $\lambda(t, x, m)$ and $\lambda_1(t, x) f(m)$. Assunção and Maia (2007) then derived the score test statistic T based on the log likelihood of the model in (8):

$$T = \sum_i \frac{\hat{\lambda}(t_i, x_i, m_i)}{\tilde{\lambda}(t_i, x_i, m_i)} - \int_{\mathcal{X}} \left[\hat{\lambda}(t, x, m) - \tilde{\lambda}(t, x, m)\right] \mathrm{d}t\, \mathrm{d}x\, \mathrm{d}m - n, \tag{9}$$

and showed using first-order Taylor expansions that the test statistics S_4 and S_5 is approximately equal to the score test statistic T under the null hypothesis of separability (2).

A sample comparison of $\hat{\lambda}(t, x, m)$ and $\tilde{\lambda}(t, x, m)$ is shown in Fig. 2a and b, which depict non-separable and separable kernel CIF estimates, respectively, of area burned and calendar date (i.e., date within the year) for the LACFD wildfires described in Sect. 2. One sees from both Fig. 2a and b that the estimated CIF increases from January to July and then decreases from August to December, with numerous small wildfires during the summer months and with large fires occurring mainly in Fall. Figure 2c and d highlight differences between the two CIF estimates: Fig. 2c shows that, from September to December, values of the non-separable intensity estimate $\hat{\lambda}(t, x, m)$ exceed those of the separable estimate $\tilde{\lambda}(t, x, m)$ for values of area burned at least 1.0 $\log \mathrm{km}^2$. Similarly, Fig. 2d shows that for areas smaller than 1.0 $\log \mathrm{km}^2$, the separable estimate is higher in June, July, August, November, and December. Figure 3 shows that these differences shown in Fig. 2c and d are statistically significant: the estimated p-value of S_3 using 400 simulations is 0, suggesting that a separable model for area burned and season might not be reasonable for the LACFD wildfires.

4 Separability of spatial–temporal marked point processes with covariates

Suppose now that the CIF for a spatial–temporal marked point process N depends not only on the time, location, and mark in question but also on space–time-varying covariates, i.e., external variables that may influence the rate of occurrence of points. For instance, in modeling the occurrence of wildfires, weather variables such as temperature and wind speed may have a substantial impact on rates of wildfire incidence (see e.g. Schroeder et al. 1964; Keeley and Fotheringham 2003). Let $c_1(t, x), \ldots, c_k(t, x)$ be k such covariates. The process N is completely separable, i.e., the spatial–temporal

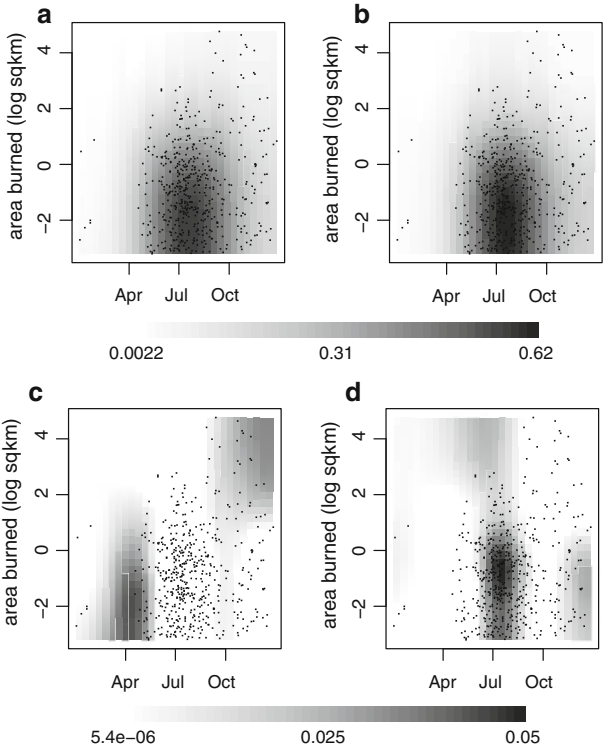

Fig. 2 CIF estimates of area burned and date: **a** non-separable kernel CIF estimate $\hat{\lambda}$, **b** separable kernel CIF estimate $\tilde{\lambda}$, **c** positive difference of $\hat{\lambda} - \tilde{\lambda}$, **d** positive difference of $\tilde{\lambda} - \hat{\lambda}$. Points on the plots indicate the 513 wildfires of at least 0.0405 km^2 (10 acres) in the LACFD dataset between January 1976 and December 2000

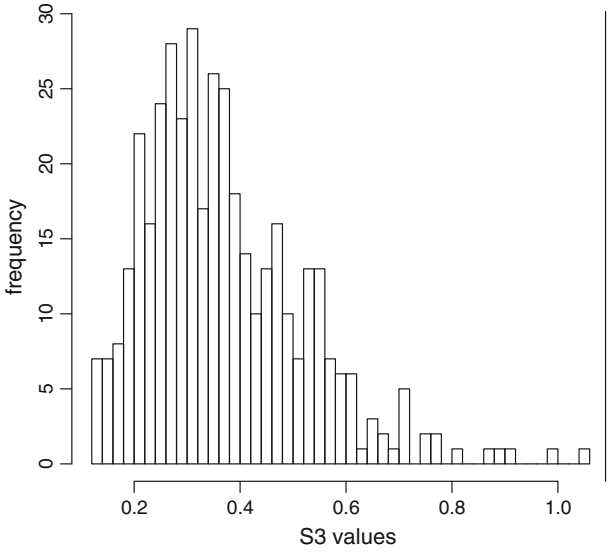

Fig. 3 Histogram of 400 simulated S_3. The *vertical line* indicates the value of S_3 calculated from the data

coordinates, the mark, and the k covariates are all separable from each other, if the CIF can be expressed as

$$\lambda[t, x, m; c_1(t, x), \ldots, c_k(t, x)] = \lambda_1(t, x) f_0(m) f_1[c_1(t, x)] \cdots f_k[c_k(t, x)], \quad (10)$$

where λ_1 is a nonnegative predictable process and $f_0, f_1, \ldots, f_k$ are fixed nonnegative functions. Complete separability (10) is especially convenient in modeling point processes with covariates, since in such cases one may readily inspect and model the influence of the covariate on the intensity of the point process simply by inspecting this covariate individually, and under quite general conditions, the parameters governing each component of the model may be consistently estimated individually by maximum likelihood (Schoenberg 2006). However, in some applications, it may be unreasonable to assume that all covariates are separable from each other.

4.1 Estimation of relative hazard for covariates

While Schoenberg (2004) suggests ordinary kernel smoothing marginal densities of the process with respect to the coordinates of the process, it is clear that different approaches are necessary in examining the separability of a spatial–temporal marked point process model with covariates. In particular, if the distribution of a covariate is non-uniform, as is typically the case for the meteorological variables recorded by the RAWS in Los Angeles County, then a simple kernel regression estimate of $f_j[c_j(t, x)]$ will be substantially biased, and an adjustment must be made.

For instance, consider a purely spatial Poisson process on a portion $\mathcal{X}$ of the plane, with intensity $\lambda(x) = \exp\{\beta_0 + \beta_1|x| + \beta_2 Z(x)\}$, where $Z(x, y)$ is a covariate, $|x|$ denotes the Euclidean norm of x, and $\beta_2 > 0$. Suppose that $Z(x, y)$ is nearly constant everywhere except at a select few sparse locations where Z attains values much higher than elsewhere in the region. Then kernel regression will yield a very poor estimate of the relationship between Z and λ; in such cases the simple kernel estimate $\tilde{\lambda}(x) = \int_{\mathcal{X}} K_1(Z(x))\mathrm{d}N(x)$ will substantially underestimate the slope in the log-linear relationship between Z and λ. The problem is that in the simple kernel estimate $\tilde{\lambda}(x)$, the large values of Z are observed much less frequently and thus are essentially counter-balanced by this lack of frequency in the product $K_1(Z(x))dN(x)$. This problem is well-known and is the basis for the Nadaraya–Watson estimator (see e.g. Silverman 1986).

To illustrate this problem further, the dotted curve in Fig. 4 shows the kernel estimate of daily 1 p.m. temperature averaged across RAWS in Los Angeles County. The estimate is slightly left-skewed and obviously non-uniform, with a peak near 21°C. The dashed curve in Fig. 4 shows a simple kernel smoothing of temperatures on days of wildfire occurrences in Los Angeles County. This is clearly a biased estimate of $f_j[c_j(t, x)]$: for instance, the estimated contribution to the CIF when temperatures are 30°C is higher than that of 40°C, since the estimate at 30°C is obtained simply by examining the number of fires occurring on days when the temperature is approximately 30°C, ignoring the fact that such days are far more commonly observed than days with 40°C. Instead, an adjustment analogous to the Nadaraya–Watson estimator

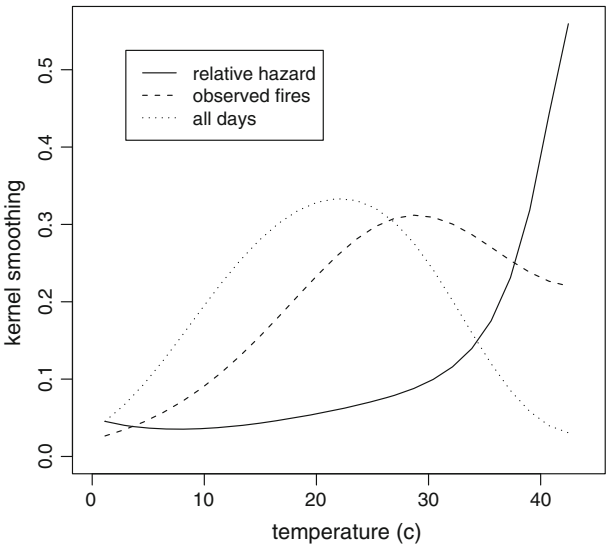

Fig. 4 Kernel estimate of temperature observed in Los Angeles County from January 1976 to December 2000. The *dashed curve* indicates the kernel estimate of temperature observed on days of wildfire occurrences (bandwidth = 9.06°C). The *dotted curve* indicates the kernel estimate of temperature observed all days (bandwidth = 5.2°C). The *solid curve* indicates the relative hazard, the expected number of wildfire occurrence per day of the corresponding temperature. Values of the *dashed curve* as well as the *dotted curve* have been modified to be included in the same figure

may be used (Silverman 1986): that is, in the case of a time-varying covariate, one may compute a quantity of the form

$$\hat{g}_j(c) = \frac{\int_{\mathcal{X}} K_1(c-v)dN(t,x,m;v)}{\int_R K_1(c-v)\mathrm{d}N(v)} = \frac{\sum_{i=1}^{n_1} K_1(c-v_i)}{\sum_{i=1}^{n_2} K_1(c-v_i)}, \tag{11}$$

where the summation in the numerator of Eq. (11) is taken over all n_1 *observed points* (fires), where the covariate (e.g., temperature) values on the days corresponding to those fires are denoted v_i, and the summation in the denominator is taken over all n_2 *days* on which the covariate is observed.

The function $g_j(c)$ may be interpreted as the proportional increase in overall wildfire hazard, or *relative hazard*, corresponding to days when the jth covariate achieves a value of c. As with kernel CIF estimates, it is natural to desire a version of f_j scaled to integrate to n_1, so a very natural estimate of the contribution $f(c)$ associated with a covariate of value c is given by

$$\hat{f}_j(c) = \frac{\hat{g}_j(c)\cdot n_1}{\int_R \hat{g}_j(v)\mathrm{d}v}. \tag{12}$$

In the case of wildfires and temperature, for instance, $\hat{f}_j(21.0)$ represents an estimate of the expected number of wildfire occurrences per day when the temperature on the given day is 21.0°C.

The solid curve in Fig. 4 shows the relative hazard $\hat{f}_j(c)$, in expected number of wildfire occurrences per day, as a function of temperature. Figure 4 suggests that the relative hazard reaches the lowest value of 0.035 wildfires per day when temperatures are 8.03°C and reaches its highest value of 0.56 wildfires per day when temperatures are 42.5°C. The relative hazard appears to increase gradually as temperature increases from 10 to 30°C, and increases rapidly for temperatures above 30°C. Note that the ordinary kernel estimate indicated by the dashed line in Fig. 4 would greatly underestimate the contribution to wildfire hazard of temperatures above 36°C, since these days occur infrequently.

4.2 Testing the separability of a covariate and a mark

One may inspect the separability of a covariate with respect to a mark of the process. Let $m(t, x) = m'$ be a real-valued mark and $c_j(t, x) = v$ be a real-valued covariate. Under the null hypothesis of complete separability (10), the CIF can be expressed as

$$\begin{aligned}\lambda(t, x, m; c_1, \ldots, c_k) &= \lambda_j(m; c_j)\lambda_{-j}(t, x, c_1, \ldots, c_{j-1}, c_{j+1}, \ldots, c_k) \\ &= f_0(m) f_j(c_j) h(t, x, c_1, \ldots, c_{j-1}, c_{j+1}, \ldots, c_k). \end{aligned} \tag{13}$$

One may compare the non-separable kernel estimate $\hat{g}_j(m; c)$:

$$\begin{aligned}\hat{g}_j(m; c) &= \frac{\int_{\mathcal{X}} K_2(m - m', c - v)\mathrm{d}N(t, x, m'; v)}{\int_R K_1(c - v)\mathrm{d}N(v)} \\ &= \frac{\sum_{i=1}^{n_1} K_1(m - m'_i)K_1(c - v_i)}{\sum_{i=1}^{n_2} K_1(c - v_i)}, \end{aligned} \tag{14}$$

with the separable kernel estimate $\tilde{g}_j(m; c)$:

$$\begin{aligned}\tilde{g}_j(m; c) &= \frac{\int_{\mathcal{X}} K_1(m - m')\mathrm{d}N(t, x, m'; c) \int_{\mathcal{X}} K_1(c - v)\mathrm{d}N(t, x, m; v)}{n_1 \cdot \int_R K_1(c - v)\mathrm{d}N(v)} \\ &= \frac{\sum_{i=1}^{n_1} K_1(m - m'_i) \sum_{i=1}^{n_1} K_1(c - v_i)}{n_1 \cdot \sum_{i=1}^{n_2} K_1(c - v_i)}. \end{aligned} \tag{15}$$

Consider the non-separable CIF estimate $\hat{\lambda}_j(m; c)$ in analogy with Eq. (12),

$$\hat{\lambda}_j(m; c) = \frac{\hat{g}_j(m; c) \cdot n_1}{\int_R \int_R \hat{g}_j(m'; v)\mathrm{d}m'\,\mathrm{d}v}. \tag{16}$$

Similarly, one may obtain a rescaled version of the separable CIF estimate $\tilde{\lambda}_j(m; c)$,

$$\tilde{\lambda}_j(m; c) = \frac{\tilde{g}_j(m; c) \cdot n_1}{\int_R \int_R \tilde{g}_j(m'; v) \mathrm{d}m' \, \mathrm{d}v}. \tag{17}$$

Thus $\hat{\lambda}_j(m; c)$ and $\tilde{\lambda}_j(m; c)$ represent the expected number of wildfire occurrences of size m on days when the jth covariate achieves a value of c. The CIF is estimated by relative hazard instead of the naive kernel estimate because the distribution of the covariate may be non-uniform.

To test the separability of the mark and the covariate c_j, one may compute any of the comparison test statistics listed in Sect. 3.3. For instance, since it is more sensitive than S_3 to the difference between $\hat{\lambda}_j(m'; v)$ and $\tilde{\lambda}_j(m'; v)$ at each observed point, one may consider the mean squared difference between the marginal CIF estimates $\hat{\lambda}_j(m; c)$ and $\tilde{\lambda}_j(m; c)$ at all points,

$$S_5 = \frac{1}{n_1} \sum_{i=1}^{n_1} \left[\hat{\lambda}_j(m_i'; v_i) - \tilde{\lambda}_j(m_i'; v_i) \right]^2. \tag{18}$$

As an illustration, Fig. 5a and b show the non-separable and separable marginal CIF estimates, respectively, for area burned and temperature. The CIF is extremely low when temperature is between 8 and 13°C, and increases steadily as temperatures increase, though there is a cluster of small wildfires when temperatures are between 24 and 32°C. The differences between the two estimates are highlighted in Fig. 5c and d. The non-separable relative hazard estimate is a bit higher than the separable estimate when temperatures are above 36°C, and the reverse appears to be the case when temperatures are between 22 and 36°C. The estimated p-value for S_5, using 400 simulations of separable processes with intensity corresponding to that in Fig. 5b, is exactly 0.05, which suggests that area burned and temperature might be nearly separable, but the differences between Fig. 5a and b are borderline statistically significant. Note that for both the non-separable and separable estimates, the relative hazards are not necessarily highest for temperatures where most of the wildfires are observed. This is because temperatures are not uniformly distributed: as described in Sect. 4.1, since the highest temperatures (such as those above 35°C are rarely observed, it is possible for the relative hazard to be highest at temperatures where relatively few wildfires actually occur.

4.3 Testing the separability of a covariate and the spatial–temporal coordinates

One may inspect the separability of a covariate with respect to one or more spatial–temporal coordinates of the process using essentially the identical method as that described in Sect. 4.2. If complete separability (10) holds, the CIF can be expressed as

$$\begin{aligned} \lambda(t, x, m; c_1, \ldots, c_k) &= \lambda_j(t, x; c_j) \lambda_{-j}(m, c_1, \ldots, c_{j-1}, c_{j+1}, \ldots, c_k) \\ &= \lambda_1(t, x) f_j(c_j) h(m, c_1, \ldots, c_{j-1}, c_{j+1}, \ldots, c_k). \end{aligned} \tag{19}$$

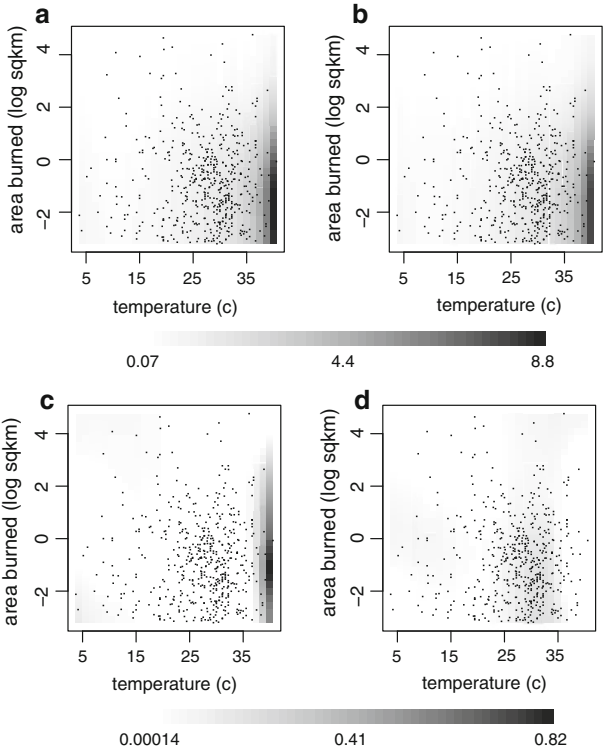

Fig. 5 CIF estimates of area burned and temperature: **a** non-separable kernel CIF estimate $\hat{\lambda}$, **b** separable kernel CIF estimate $\tilde{\lambda}$, **c** positive difference of $\hat{\lambda} - \tilde{\lambda}$, **d** positive difference of $\tilde{\lambda} - \hat{\lambda}$. Points on the plots correspond to observed wildfires

Specifically, one may compare the non-separable kernel estimate

$$\hat{g}_j(t, x; c) = \frac{\sum_{i=1}^{n_1} K_{n+1}(t - u_i, x - y_i) K_1(c - v_i)}{\sum_{i=1}^{n_2} K_1(c - v_i)} \tag{20}$$

with the separable estimate

$$\tilde{g}_j(t, x; c) = \frac{\sum_{i=1}^{n_1} K_{n+1}(t - u_i, x - y_i) \sum_{i=1}^{n_1} K_1(c - v_i)}{n_1 \cdot \sum_{i=1}^{n_2} K_1(c - v_i)} \tag{21}$$

in order to inspect whether the marginal CIF estimates

$$\hat{\lambda}_j(t, x; c) = \frac{\hat{g}_j(t, x; c) \cdot n_1}{\int_R \int_{R^n} \int_0^T \hat{g}_j(u, y; v) \mathrm{d}u \, \mathrm{d}y \, \mathrm{d}v} \tag{22}$$

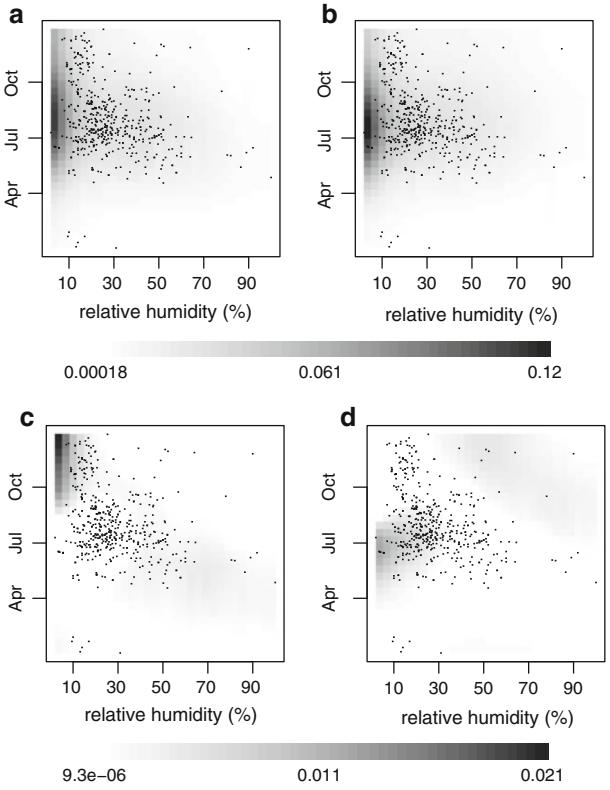

Fig. 6 CIF estimates of date and relative humidity: **a** non-separable kernel CIF estimate $\hat{\lambda}$, **b** separable kernel CIF estimate $\tilde{\lambda}$, **c** positive difference of $\hat{\lambda} - \tilde{\lambda}$, **d** positive difference of $\tilde{\lambda} - \hat{\lambda}$. Points on the plots correspond to observed wildfires

and

$$\tilde{\lambda}_j(t, x; c) = \frac{\tilde{g}_j(t, x; c) \cdot n_1}{\int_R \int_{R^n} \int_0^T \tilde{g}_j(u, y; v) \mathrm{d}u \, \mathrm{d}y \, \mathrm{d}v} \tag{23}$$

appear to be separable. Again, the CIF is estimated by relative hazard instead of the naive kernel estimate because the distribution of the covariate c_j may be non-uniform. As for the case of testing the separability of a covariate and the spatial–temporal coordinates, we suggest the mean squared difference between $\hat{\lambda}_j(t, x; c)$ and $\tilde{\lambda}_j(t, x; c)$ at all points,

$$S_5 = \frac{1}{n_1} \sum_{i=1}^{n_1} \left[\hat{\lambda}_j(u_i, y_i; v_i) - \tilde{\lambda}_j(u_i, y_i; v_i) \right]^2. \tag{24}$$

Figure 6 shows the non-separable and separable CIF estimates for date and relative humidity. The estimated CIF gradually decreases as relative humidity increases, and

is extremely low when relative humidity is above 90%. When relative humidity is below 10%, one sees that the estimated CIF appears to increase from April to July and decrease from August to December, while wildfire incidence in January and February is minimal. Figure 6c shows that the non-separable estimate in Fig. 6a is a bit higher than the separable estimate in Fig. 6b from mid-August to December when relative humidity is below 20%. Figure 6d shows that the separable estimate is higher from April to August when relative humidity is below 35%. The estimated p-value corresponding to the statistic S_5 using 400 simulations of separable processes with CIF shown in Fig. 6b is 0, which suggests that the differences between Fig. 6a and b are significant and thus date and relative humidity appear to be significantly non-separable. Much of the reason for this significance seems to be due to the fact that, for the simulated separable point processes, each simulation had more wildfires, compared to the actual dataset, from April to August when relative humidity was below 35%, and generally fewer wildfires elsewhere. Therefore, the difference between $\hat{\lambda}$ and $\tilde{\lambda}$ became smaller for each simulation, resulting in a smaller value of S_5 for the simulations, compared to that for the data.

4.4 Testing the separability of two covariates

One might inspect the separability of a covariate with respect to another covariate influencing the same point process. Let $c_i(t, x) = v_1$ and $c_j(t, x) = v_2$ be real-valued covariates. Under the null hypothesis of complete separability (10), the CIF can be expressed as

$$\begin{aligned}\lambda(t, x, m; c_1, \ldots, c_k) &= \lambda_{ij}(c_i, c_j)\lambda_{-ij}(t, x, m, c_{-ij}) \\ &= f_i(c_i) f_j(c_j) h(t, x, m, c_{-ij})\end{aligned} \tag{25}$$

for $i \neq j$, where c_{-ij} indicates all covariates except c_i and c_j. Unlike the methods described in Sects. 4.2 and 4.3, a slight modification is typically necessary in testing the separability of the CIF with respect to two covariates.

Nonseparable and separable kernel estimates of the hazard associated with covariates c_i and c_j may readily be obtained via

$$\hat{g}_{ij}(c_1, c_2) = \frac{\sum_{i=1}^{n_1} K_1(c_1 - v_{1i}) K_1(c_2 - v_{2i})}{\sum_{i=1}^{n_2} K_1(c_1 - v_{1i}) K_1(c_2 - v_{2i})} \tag{26}$$

and

$$\tilde{g}_{ij}(c_1, c_2) = \frac{n_2 \cdot \sum_{i=1}^{n_1} K_1(c_1 - v_{1i}) \sum_{i=1}^{n_1} K_1(c_2 - v_{2i})}{n_1 \cdot \sum_{i=1}^{n_2} K_1(c_1 - v_{1i}) \sum_{i=1}^{n_2} K_1(c_2 - v_{2i})}, \tag{27}$$

respectively. Consider the denominator of Eq. (26), which indicates the bivariate kernel smoothing of the total number of days when the covariates achieved corresponding values near c_1 and c_2. If such days rarely occur, then the bivariate kernel estimate will assume a very low value, and in practice may even achieve values indistinguishable

from zero using standard computer packages. For instance, days with temperature near 34°C and relative humidity near 93% did not occur in Los Angeles County during the time range considered; thus, the bivariate kernel estimate of such days has a value of essentially 0 wildfires per day. In order to restrict attention only to conditional intensities that may be stably estimated, one may wish to omit values of c_1 and c_2 corresponding to extremely low values of the denominator of Eq. (26) from the analysis. Define the cut-off value λ_c as $\lambda_c = \min \hat{g}_{ij}(v_1, v_2)$, i.e., the cut-off value should be small enough to include all n_2 days but be sufficiently large so that it is insensitive to the extremely low values of the denominator of Eqs. (26) and 27. Note that the same λ_c should be applied to both Eqs. (26) and (27) for comparison. One may compute a non-separable CIF estimate $\hat{\lambda}_{ij}(c_1, c_2)$ via:

$$\hat{\lambda}_{ij}(c_1, c_2) = \frac{\hat{h}_{ij}(v_1, v_2) \cdot n_1}{\int_R \int_R \hat{h}_{ij}(v_1, v_2) \mathrm{d}v_1 v_2}, \tag{28}$$

where $\hat{h}_{ij}(c_1, c_2) = \hat{g}_{ij}(c_1, c_2) \cdot 1\{\hat{g}_{ij}(c_1, c_2) \geq \lambda_c\}$, as well as the separable CIF estimate

$$\tilde{\lambda}_{ij}(c_1, c_2) = \frac{\tilde{h}_{ij}(c_1, c_2) \cdot n_1}{\int_R \int_R \tilde{h}_{ij}(v_1, v_2) \mathrm{d}v_1 \, \mathrm{d}v_2}, \tag{29}$$

where $\tilde{h}_{ij}(c_1, c_2) = \tilde{g}_{ij}(c_1, c_2) \cdot 1\{\tilde{g}_{ij}(c_1, c_2) \geq \lambda_c\}$.

To test the separability of c_i and c_j, we suggest the mean squared difference between $\hat{\lambda}_{ij}(c_1, c_2)$ and $\tilde{\lambda}_{ij}(c_1, c_2)$ at all points,

$$S_5 = \frac{1}{n_2} \sum_{i=1}^{n_2} \left[\hat{\lambda}_{ij}(v_{1i}, v_{2i}) - \tilde{\lambda}_{ij}(v_{1i}, v_{2i}) \right]^2, \tag{30}$$

because one expects the resulting estimate to be especially powerful at detecting departures from separability at values corresponding to actual observed values of the covariates.

Figure 7 shows non-separable and separable CIF estimates of temperature and relative humidity, corresponding to Eqs. (28) and (29), respectively, using a cut-off value of 0.1179. The shading of each pixel indicates an estimate of the expected number of wildfire occurrences per day of the corresponding temperature and relative humidity. Figure 7c shows that the non-separable CIF estimate is a bit higher than the separable CIF estimate at the boundaries. The separable CIF estimate in Fig. 7d is slightly higher when temperatures are above 27°C and when relative humidity is below 40%. The estimated p-value for S_5, based on 100 simulations of separable processes with CIF shown in Fig. 7b, is 0, which suggests that the differences between Fig. 7a and b are significant and thus temperature and relative humidity appear to be significantly non-separable.

Figure 8 shows non-separable and separable CIF estimates of relative humidity and wind speed using $\lambda_c = 0.0445$. The CIF estimates are high when relative humidity is below 10% and when wind speed is above 40 km h^{-1}; the CIF is low when relative

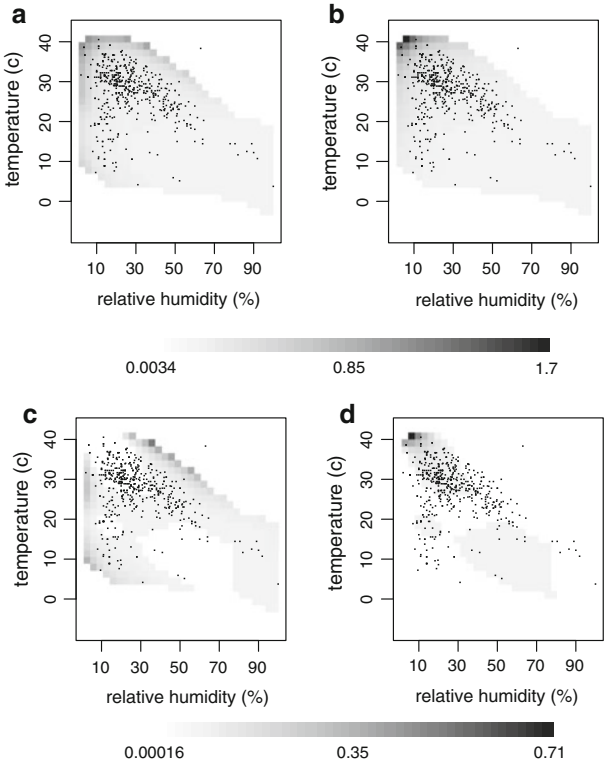

Fig. 7 CIF estimates of temperature and relative humidity: **a** non-separable kernel CIF estimate $\hat{\lambda}$, **b** separable kernel CIF estimate $\tilde{\lambda}$, **c** positive difference of $\hat{\lambda} - \tilde{\lambda}$, **d** positive difference of $\tilde{\lambda} - \hat{\lambda}$

humidity is above 70%. The non-separable CIF estimate is higher than the separable CIF estimate when relative humidity is below 10% and when wind speed is below 30 km h^{-1}. Figure 8d shows that the separable CIF estimate is higher when wind speed is greater than 30 km h^{-1} and when relative humidity is below 40%. The estimated p-value for S_5, using 100 simulations of separable processes with CIF shown in Fig. 8b, is 0.17, which suggests that the differences shown in Fig. 8c and d are statistically insignificant and thus relative humidity and wind speed might be separable.

Precipitation is a special covariate. From January 1976 to December 2000, only 15 out of 513 wildfires occurred on days with non-zero recorded precipitation. In fact, although RAWS are very sensitive even to very small quantities of precipitation, records at all available RAWS stations were identically zero for 85.7% of days during this 25-year period considered. Such an overlap of identical values presents problems for standard kernel smoothing estimates, since the bandwidth estimates obtained using standard formulas are generally too small in such circumstances (Silverman 1986; Schoenberg et al. 2009). One option is simply to remove the entire cluster of overlapping values from the dataset before fitting any kernel estimates at other values. Figure 9 shows the non-separable and separable CIF estimates of relative humidity and precipitation after removing zeros from the dataset, and using $\lambda_c = 0.7076$. The 15

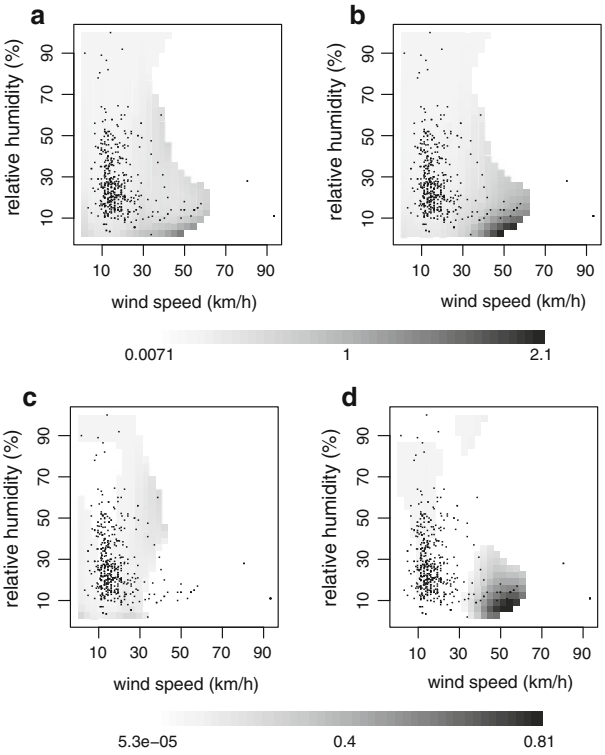

Fig. 8 CIF estimates of relative humidity and wind speed: **a** non-separable kernel CIF estimate $\hat{\lambda}$, **b** separable kernel CIF estimate $\tilde{\lambda}$, **c** positive difference of $\hat{\lambda} - \tilde{\lambda}$, **d** positive difference of $\tilde{\lambda} - \hat{\lambda}$. Points on the plots correspond to observed wildfires

points on the plot indicate the 15 wildfires occurring on days with non-zero precipitation during the time range considered. The shading of each pixel indicates the expected number of wildfire occurrences per day of the corresponding relative humidity and precipitation. The estimated p-value for S_5, based on 100 simulations of separable processes with CIF shown in Fig. 9b, is 0.62, suggesting that any departures from separability, for the covariates relative humidity and precipitation, do not appear to be statistically significant.

5 Non-separability and partitioning

In cases where the separability hypothesis is seriously violated for a pair of coordinates, Schoenberg (2006) suggested partitioning one of the coordinates and fitting a separable model within each segment of the partition. For example, as illustrated in Sect. 3.3, area burned and date appear to be highly non-separable for the Los Angeles County wildfire data; the test statistic S_3 for area burned and date has a value of 1.09, and a corresponding p-value, using 400 simulations, of 0. This suggests that the distribution of burn area changes significantly with date.

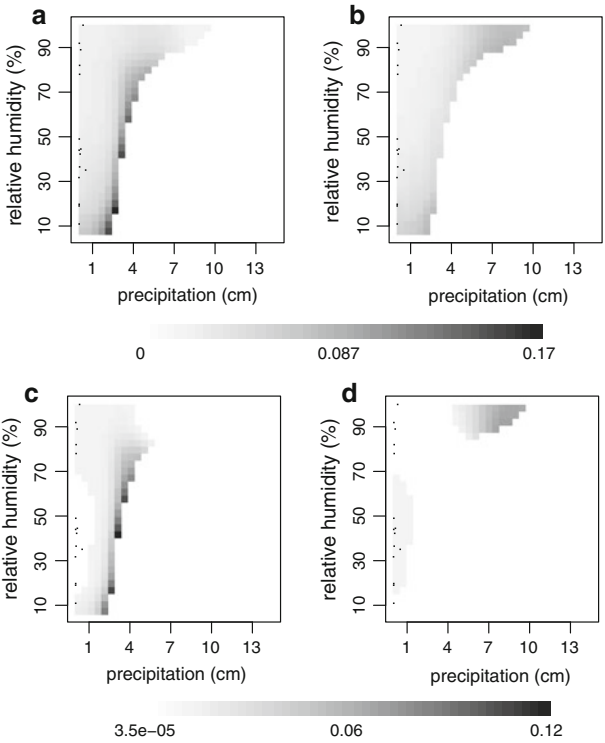

Fig. 9 CIF estimates of relative humidity and precipitation: **a** non-separable kernel CIF estimate $\hat{\lambda}$, **b** separable kernel CIF estimate $\tilde{\lambda}$, **c** positive difference of $\hat{\lambda} - \tilde{\lambda}$, **d** positive difference of $\tilde{\lambda} - \hat{\lambda}$. Points on the plots correspond to observed wildfires

Table 1 p-Values for the separability of area burned and date, for data within different seasons

	Season 1	Season 2	Season 3
Area burned	0.3525	0.3075	0.7675

One way to approach modeling in such non-separable cases is to consider dividing the dataset according to different seasons as suggested in Schoenberg (2006). For instance, one may consider the following three seasonal divisions: (1) April 1 to June 30; (2) July 1 to September 30; and (3) October 1 to March 31. Table 1 shows the p-values corresponding to an S_3 statistic testing the separability of area burned and date, for each season individually. The distribution of area burned appears to be relatively constant within each season. This suggests replacing (3) with a seasonal model such as

$$\lambda(t, x, m) = \lambda_1(t) f_1(x) \sum_{j=1}^{3} f_2^{(j)}(m) 1_{\{t \in T_j\}}, \tag{31}$$

where T_j, $j = 1, 2, 3$, are divisions of the observed time span corresponding to the three seasons listed above, and $f_2^{(j)}(m)$ represents the burn area density during season j. The identical adjustment can be made for the case where a pair of covariates is non-separable or when a covariate and mark are non-separable.

6 Discussion

The importance of separability in the modeling of multi-dimensional point processes cannot be overstated. Indeed, when separability may safely be assumed, not only is the derivation of a parametric form for the relationship between one covariate and the CIF greatly facilitated due to an effective reduction (to unity) of the dimensionality of the process, but in addition the parameters governing each component of the model may be estimated individually with consistency (Schoenberg 2006). Hence the separability tests of Schoenberg (2004), which are here extended to the case of multiple covariates, should be used in the construction of models for marked multivariate point processes with covariates, such as the case of estimating the hazard of Los Angeles County wildfires as a function of daily weather variables.

For illustrative purposes, we have focused the attention here mainly on pairs of covariates (or pairs consisting of a covariate and a mark or a covariate and a spatial–temporal coordinate of the process) demonstrating significant departures from separability, in order to highlight the separability tests proposed. However, for most such pairs, no significant departures from separability were present. The extent to which date and burn area appear to be non-separable is not surprising since the variable distribution of wildfire burn area in Los Angeles County across different seasons is well-known (see e.g. Keeley 2002). It is somewhat remarkable, however, how dramatically the significance of this departure from separability is removed when the dataset is partitioned according to season. It is also a bit surprising that temperature and relative humidity are so similarly non-separable. As noted in Sect. 4.4, this seems largely due to differences between estimates near the boundaries, particularly on days of extremely low relative humidity. Note that since all the separability tests used here are non-parametric, such departures from non-separability cannot be attributed to the lack of fit of a model used in the estimation of the contribution from a particular covariate. The investigation of the power and other statistical properties of the tests discussed here, as well as the application of separability tests to other wildfire datasets and to other marked point process data with covariates, are important topics for future work.

Acknowledgments Thanks to James Woods, the LACDPW and LACFD (especially Mike Takeshita and Frank Vidales) for their generosity in sharing data, and to Haiyong Xu and Jamie Pompa for valuable suggestions.

References

Andrews, P. L., Bradshaw, L. S. (1997). FIRES: Fire Information Retrieval and Evaluation System: A program for fire danger rating analysis. Gen. Tech. Rep. INT-GTR-367. Ogden, UT: US Department of Agriculture, Forest Service, Intermountain Research Station. p. 64

Assunção, R., Maia, A. (2007). A note on testing separability in spatial–temporal-marked point processes. *Biometrics, 63*, 290–294.

Bowman, A. W., Azzalini, A. (1997). *Applied smoothing techniques for data analysis: The kernel approach with S-Plus illustrations*. Oxford: Clarendon Press.

Bradshaw, L. S., Deeming, J. E., Burgan, R. E., Cohen, J. D. (1983). The 1978 national fire-danger rating system: Technical documentation. United States Department of Agriculture Forest Service General Technical Report INT-169. Ogden, UT: Intermountain Forest and Range Experiment Station. p. 46.

Cressie, N. A. (1993). *Statistics for spatial data* (revised ed.). New York: Wiley.

Daley, D., Vere-Jones, D. (2003). *An introduction to the theory of point processes* (Vol. I, 2nd ed.). New York: Springer.

Kagan, Y. Y., Jackson, D. D. (1994). Long-term probabilistic forecasting of earthquakes. *Journal of Geophysical Research, 99*, 13685–13700.

Keeley, J. (2002). Fire management of California shrubland landscapes. *Environmental Management, 29*, 395–408.

Keeley, J. E., Fotheringham, C. J. (2003). Impact of past, present, and future fire regimes on North American Mediterranean shrublands. In T. T. Veblen, W. L. Baker, G. Montenegro, T. W. Swetnam (Eds.), *Fire and climatic change in temperate ecosystems of the Western Americas* (pp. 218–262). New York: Springer

Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical Association, 83*(401), 9–27.

Ogata, Y. (1998). Space–time point-process models for earthquake occurrences. *Annals of the Institute of Statistical Mathematics, 50*(2), 379–402.

Peng, R. D., Schoenberg, F. P., Woods, J. (2005). A space–time conditional intensity model for evaluating a wildfire hazard index. *Journal of the American Statistical Association, 100*, 26–35.

Pyne, S. J., Andrews, P. L., Laven, R. D. (1996). *Introduction to wildland fire* (2nd ed.). New York: Wiley.

Schoenberg, F. P. (2004). Testing separability in multi-dimensional point processes. *Biometrics, 60*, 471–481.

Schoenberg, F. P. (2006). A note on the separability of multidimensional point processes with covariates. UCLA Preprint Series, No. 496.

Schoenberg, F. P., Brillinger, D. R., Guttorp, P. M. (2002). Point processes, spatial–temporal. In A. El-Schaarawi, W. Piegorsch (Eds.), *Encyclopedia of environmetrics* (Vol. 3, pp. 1573–1577). New York: Wiley.

Schoengerg, F. P., Peng, R., Woods, J. (2003). On the distribution of wildfire sizes. *Environmetrics, 14*, 583–592.

Schoenberg, F. P., Chang, C., Keeley, J. E., Pompa, J., Woods, J., Xu, H. (2007). A critical assessment of the Burning Index in Los Angeles County, California. *International Journal of Wildland Fire, 16*, 473–483.

Schoenberg, F. P., Pompa, J. L., Chang, C. (2009). A note on non-parametric and semi-parametric modeling of wildfire hazard in Los Angeles County, California. *Journal of Environmental and Ecological Statistics, 16*, in press.

Schroeder, M. J., Glovinsky, M., Hendricks, V., Hood, F., Hull, M., Jacobson, H., Kirkpatrick, R., Krueger, D., Mallory, L., Oertel, A., Reese, R., Sergius, L., Syverson, C. (1964). *Synoptic weather types associated with critical fire weather*. Washington, DC: Institute for Applied Technology, National Bureau of Standards, US Department of Commerce.

Sheather, S. J., Jones, M. C. (1991). A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society, Series B, 53*, 683–690.

Silverman, B. W. (1986). *Density estimation for statistics and data analysis*. London: Chapman and Hall.

Stone, C. J. (1984). An asymptotically optimal window selection rule for kernel density estimates. *The Annals of Statistics, 12*, 1285–1297.

Venables, W. N., Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.). New York: Springer.

Warren, J. R., Vance, D. L. (1981). Remote automatic weather station for resource and fire management agencies. Technical Report INT-116. Ogden, UT: USDA Forest Service, Intermountain Forest and Range Experiment Station.